



## UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/634,145	08/04/2003	Chew Kiat Heng	NAA 0018 PA/41049.20	5097
23368	7590	10/27/2009		
DINSMORE & SHOHL LLP ONE DAYTON CENTRE, ONE SOUTH MAIN STREET SUITE 1300 DAYTON, OH 45402-2023			EXAMINER	
			WHALEY, PABLO S	
			ART UNIT	PAPER NUMBER
			1631	
			MAIL DATE	DELIVERY MODE
			10/27/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 10/634,145	<b>Applicant(s)</b> HENG ET AL.
	<b>Examiner</b> PABLO WHALEY	<b>Art Unit</b> 1631

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED. (35 U.S.C. § 133).

Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

1) Responsive to communication(s) filed on 13 August 2009.

2a) This action is FINAL.      2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

4) Claim(s) 1-15, 17-26 and 28-30 is/are pending in the application.

4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.

5) Claim(s) \_\_\_\_\_ is/are allowed.

6) Claim(s) 1-15, 17-26, and 28-30 is/are rejected.

7) Claim(s) \_\_\_\_\_ is/are objected to.

8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All    b) Some \* c) None of:  
 1. Certified copies of the priority documents have been received.  
 2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

1) Notice of References Cited (PTO-892)  
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)  
 3) Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_

4) Interview Summary (PTO-413)  
 Paper No(s)/Mail Date. \_\_\_\_\_

5) Notice of Informal Patent Application  
 6) Other: \_\_\_\_\_

**DETAILED ACTION**

***Request For Continued Examination***

A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 08/13/2009 has been entered.

***Status of Claims***

Claims 1-15, 17-26, and 28-30 are pending and under consideration.

Claims 16 and 27 are cancelled.

***Withdrawn Rejections***

The rejection of claims 1-11, 18-23, 28, 29, and 30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, and Walters is withdrawn in view of applicant's arguments filed 08/13/2009 that Tibshirani in does not teach weights determined with a constraint that weights associated with sets of data having like genetic data are the same, or a fitting process wherein a calculation of deviates includes a weighting of weights.

The rejection of claims 1-15 and 17- 30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, Walters, and Lazzeroni is withdrawn in view of applicant's arguments filed 08/13/2009 for reasons set forth above.

The rejection of claims 1-11, 18-26, and 28-30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, Walters, and Nelson is withdrawn in view of applicant's arguments filed 08/13/2009 for reasons set forth above.

*Claim rejections - 35 USC § 112, 2<sup>nd</sup> Paragraph*

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-15, 17-26, and 28-30 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Claims that depend directly or indirectly from claims 1, 4, and 21 are also rejected due to said dependence.

Claims 1 and 21 recite the step of model fitting comprising "calculating a sum of weighted deviates for all of said sets, wherein each deviate is weighted in said sum by a weight associated with, and indicating a statistical significance of, that set for which said each deviate has been calculated, and wherein the weights used to weight said deviates are determined with a constraint that said weights associated with sets of said data having like genetic data are the same." This limitation is confusing for the following reasons.

It is unclear what limitation of the claimed method is intended by "wherein the weights used to weight said deviates are determined with a constraint that said weights associated with sets of said data having like genetic data are the same." This could be interpreted as an active method step (e.g. determining weights using a constraint) or simply as a further limitation of the weights. Clarification is requested via clearer claim language. It is noted that the nature of the data (e.g. weights associated with data), per se, has no restrictive effect on the claimed method. Therefore, the Examiner has broadly interpreted the claims for purposes of applying prior art.

It is also unclear what limitation of the claimed method is intended by "weights associated with sets of said data having like genetic data are the same." One interpretation of this limitation is that the

weights are the same. Another interpretation is that the sets of data (having like genetic data) are the same. Clarification is requested via clearer claim language.

Claim 4 (line 4) recites the limitation "data having like genetic data like genetic data." It is unclear what limitation of said data is intended by repeating the phrase "like genetic data" twice. Clarification is requested via clearer claim language.

Claim 4 (last two lines) recites the limitation "whereby the group weight is the corresponding weight for each set of data." The term "corresponding weight" implies that weights for each group are related by some existing criteria. However, the specification does not provide a standard or criteria for determining corresponding weights such that one of ordinary skill in the art would know the metes and bounds of this limitation, as claimed. Clarification is requested via clearer claim language.

#### *Claim Rejections - 35 USC § 103*

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966),

that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

Claims 1-6, 9-11, 13-15, 17-21, and 28-30 are rejected under 35 U.S.C. 103(a) as being made obvious by Parzen (*Biometrics*, 1999, Vol. 55, p.580-584), in view of Shattuck-Eidens et al. (*JAMA*, 1997, Vol. 278, No. 15, p. 1242-1250), and in view of Cleveland (*Journal of the American Statistical Association*, 1979, Vol. 74, No. 368, p.829-836).

The claims are drawn to a computer-implemented method of determining a statistical model for predicting disease risk for a member of a population. The claims required steps for collecting genetic and non-genetic data, and an indicator of disease status. The claims require storing a candidate model for calculating risk as a function of non-genetic data and plurality of parameters. The claims require steps for optimizing these parameters by fitting, wherein the fitting comprises a plurality of steps including calculating a deviate of a predicted risk from said indicator of disease status, wherein risk is predicted using the model, calculating a sum of weighted deviates for all data sets, wherein each deviate is weighted by a weight determined with the constraint that the weights associated with data having like genetic data are the same, and minimizing the sum of weighted deviates to obtain optimized parameters. The result of the claims is that risk is calculated using the model with optimized parameters and non-genetic data. For purposes of examination, genetic data and non-genetic data are interpreted as genetic and non-genetic risk factors in view of the specification [0030].

Parzen teaches a method for optimizing linear regression models used to predict liver disease [Abstract]. In particular, Parzen shows a Cox hazard regression model for calculating disease risk in a subject are described in full [Section 2]. The model is a linear combination of coefficients and covariates that include age, albumin, and edema data sets [Table 1, Table 2, and p.581, Col. 2], which are interpreted as non-genetic risk factors. An indicator of disease status is described, N(t), which fluctuates between 1 and 0 over time based on patient risk [p.581, Col. 2]. Parzen calculates partial risk estimates using a Cox likelihood score vector wherein Z is based on a sum of weighted averages and dN is a binary variable between 1 and 0 (i.e. weight) [p.581, Col. 2, Equation 2], which is interpreted as a target function. Parzen describes an optimization procedure based on curve-fitting [Section 3]. In particular, data is partitioned into groups and group weights (I) are assigned a value of 1 or 0 [p.581, last ¶]. If the model is correctly specified, parameters for an arbitrary number of groups will take the value of zero in the Cox model [p.582, Col. 1, ¶1 and Equation 3], which shows weights associated with sets of data having like values. Subjects in the same group can also be considered similar if they have similar risks at any given time [p.581, Col. 1]. Parzen also calculates the Chi-squared distribution as an alternative measure of goodness of fit [p.582, Col. 1]. Parzen also defines a residual equation for calculating goodness of fit based the difference between observed minus expected number of failures in each region [p.582, Col. 2], and calculates the total number of failures based on the sum of the estimated expected failures. The Chi-squared distribution is interpreted as a teaching for calculating weighted deviates since it is used in the model fitting process and is based on weighted deviations in the data. Parzen shows selecting the model with a minimized goodness of fit statistic [p. 582, Col. 1 and 583, Col. 1, ¶2]. Parzen shows two different models with the same Zi number of parameters [Equations 1 and 3].

Parzen does not teach collecting genetic data sets associated with members of a population, or both genetic and non-genetic factors, as in claims 1, 17, 21, and 28.

Parzen does not specifically teach calculating a sum of weighted deviates for all data sets, wherein the weights used to weight said deviates are determined with the constraint that the weights associated with data having like genetic data are the same, as in claims 1, 21, and 28.

Parzen does not teach a plurality of weights that are weighted by an adjustment factor, as in claims 13 and 14.

Shattuck-Eidens teaches a statistical model for predicting disease risk using a plurality of genetic risk factors [See at least p.1243, Col. 3, and p.1244 and Tables 1-6]. Patients are also classified according to specific characteristics [p.1246]. This method is beneficial for predicting cancer in patients with detected genetic mutations [p.1244].

Cleveland teaches a computer-based method for optimizing models based on weighted regression. Given a linear model, Cleveland shows calculating a sum of weighted deviates [p.830, Col. 2]. In particular, the model is optimized by re-fitting the regression model using the newly calculated weighted deviate values [See steps 1-4, p. 830 and 831]. Cleveland also shows a weighting function wherein values above a certain x threshold all equal 0 [p.831, Col. 1], which suggests equal weights for certain points in a data set. Cleveland also shows an optimization process that includes a robustness weight calculation that is used to weight different weights and is based on a ratio of residuals and the median [p.831, Col. 1], which shows weights weighted by an adjustment factor. Cleveland further optimizes parameters based on error variance and linear sum of residuals [Section 6.1 and p.835, Col. 1]. This technique is beneficial for smoothing distortions in data [Section 4.4]. Cleveland shows techniques for reducing computations [Section 5.1], which inherently shows the use of computers and computer software for performing these methods.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method of Parzen by collecting genetic data sets associated with members of a

population, and both genetic and non-genetic factors, as in claims 1, 17, 21, and 28, since Shattuck-Eidens uses as model for predicting disease risk using a plurality of genetic risk factors [See at least p.1243, Col. 3, and p.1244 and Tables 1-6] with predictable results. The motivation would have been to use regression models for describing the relationship between multiple variables related to disease.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method of Parzen by calculating a sum of weighted deviates for any of the collected data sets, wherein the weights used to weight said deviates are determined with the constraint that the weights associated with data having like genetic data are the same, as in claims 1, 21, and 28, since Parzen shows that model fitting based on Chi-squared deviations is well known and since Cleveland shows model optimization by fitting parameters based on weighted deviate calculations with predictable results, as set forth above, [See steps 1-4, p. 830 and 831]. The motivation would have been to improve the disease model by finding values for the coefficients such that the regression model matches the raw data as well as possible.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method of Parzen by determining a plurality of weights that are weighted by an adjustment factor, as in claims 13 and 14, since the selection of weights in general is an arbitrary design consideration based on the nature of the data, as suggested by Cleveland, and since Shattuck-Eidens shows accounting for the influence of different groups on the regression model using weights [p.1246, Col. 2, Col. 3]. The motivation would have been to use weights to correct for possible bias in the population data.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to practice the method of Parzen by using a computer and computer software since Shattuck-Eidens and Cleveland suggests such prediction methods are designed for computers, as set forth above.

The motivation would have been to improve disease prediction using automated techniques for performing complex calculations.

Claims 7, 8, 12, and 22-26 are rejected under 35 U.S.C. 103(a) as being made obvious by Parzen (*Biometrics*, 1999, Vol. 55, p.580-584), in view of Shattuck-Eidens et al. (*JAMA*, 1997, Vol. 278, No. 15, p. 1242-1250), and in view of Cleveland (*Journal of the American Statistical Association*, 1979, Vol. 74, No. 368, p.829-836), as applied to claims 1-6, 9-11, 13-15, 17-21, and 28-30, above, and further in view of Kooperberg et al. (*Technical Report*, 1996, p.1-20) and Hu et al. (*Proceedings of the Survey Research Methods Section, ASA*, 1996, p.287-292).

Parzen, Shattuck-Eidens, and Cleveland make obvious a method for determining a model for predicting disease risk, as set forth above. Additionally, Cleveland shows a function based on a summation of weights and residual size [See at least p.830, Col. 1, Col. 2, and p.834, Col. 1]. Shattuck-Eidens also shows correlating risk factors and grouping risk factors based on clustering [p.1246 and Table 6].

Parzen, Shattuck-Eidens, and Cleveland do not teach a residual for an itch one of said data sets in said reference group that is the difference between a value of the indicator of disease status contained in said itch data set and the value of disease risk for the member associated with said itch data set, said value of disease risk calculated from said candidate model with said parameters optimized for a given set of group weights by fitting data sets in groups other than the reference group to said candidate model, as in claim 7.

Parzen, Shattuck-Eidens, and Cleveland do not teach imputing missing data, as in claims 12 and 22.

Parzen, Shattuck-Eidens, and Cleveland do not teach dividing data, and recursive division, as in claims 23, 24, 25, and 26.

Parzen, Shattuck-Eidens, and Cleveland do not teach determining if a criteria is met after dividing, said criteria is evaluated based on genetic data in each of said data sets, and regrouping when criteria is not met, as in claim 23.

Parzen, Shattuck-Eidens, and Cleveland do not teach performing division recursively on each group of a division, and wherein divisions are made dependent on data indicative of different factors, as in claims 24-26.

Methods for dividing data within predictive modeling processes are well known. In particular, Kooperberg teaches methods for selecting optimal models by dividing data into equally sized subgroups with the constraint that data not in the j-th subgroup is fitted to the model [See Section 3.2, p.6-7], which is interpreted as fitting data sets in groups other than the reference group. The best model is selected by minimizing a cross-validation loss function for data not used to fit the model [See Section 3.2, p.6-7].

Methods for imputing data with predictive modeling processes are well known. In particular, Hu shows software for imputing missing values in regression models. The software program partitions the range of regression values from the data set into subsets [p.287, Col. 2, ¶2, p.288, Col. 2, ¶2]. Weighted average values are then computed and assigned to subsets with missing data [p.287, Col. 2, ¶2]. The subsets are assumed to be homogeneous. This technique is beneficial for eliminating bias in large data sets [Section II and p.292, Col. 2].

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method made obvious by Parzen, Shattuck-Eidens, and Cleveland by determining a residual for an itch one of said data sets in said reference group that is the difference between a value of the indicator of disease status contained in said itch data set and the value of disease risk for the member

Art Unit: 1631

associated with said itch data set, said value of disease risk calculated from said candidate model with said parameters optimized for a given set of group weights by fitting data sets in groups other than the reference group to said candidate model, as in claim 7, since Parzen and Cleveland already suggest models for determining residuals between any elements in the data set, as set forth above, and since Kooperberg suggests improving models by fitting them using data sets in groups other than a reference group, as set forth above. The motivation would have been the use of cross-validation to estimate disease risk.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method made obvious by Parzen, Shattuck-Eidens, and Cleveland by imputing missing data, as in claims 12 and 22, since Shattuck-Eidens uses data that includes missing data sets [Table 5], and since Hu shows software for imputing missing values in regression models, as set forth above. The motivation would have been to eliminate bias in large data sets [Section II and p.292, Col. 2].

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to modify the method made obvious by Parzen, Shattuck-Eidens, and Cleveland by dividing data based on criteria, as in claims 23, 24, 25, and 26, since Kooperberg and Hu shows data partitioning using criteria specifically related to the data sets is well established, as shown above, and since Cleveland shows dividing data into groups data using iterative criteria can beneficially reduce the computational load on a computer [Section 5.1].

***Response to Argument***

Applicants' arguments, filed 08/13/2009, that Tibshirani does not teach weights determined with a constraint that weights associated with sets of data having like genetic data are the same, or a fitting process wherein a calculation of deviates includes a weighting of weights, have been fully considered and

Art Unit: 1631

are persuasive. The rejection of claims 1-11, 18-23, 28, 29, and 30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, and Walters is withdrawn. The rejection of claims 1-15 and 17- 30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, Walters, and Lazzeroni is withdrawn. The rejection of claims 1-11, 18-26, and 28-30 under 35 U.S.C. 103(a) as being made obvious by Tibshirani in view of Nguyen, Mariani, Walters, and Nelson is withdrawn.

---

*Conclusion*

No claim is allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Pablo Whaley whose telephone number is (571)272-4425. The examiner can normally be reached on 9:30am - 6pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marjorie Moran can be reached at 571-272-0720. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Pablo S. Whaley

Patent Examiner

Art Unit: 1631

Art Unit 1631

/PW/

/SHUBO (Joe) ZHOU/

Primary Examiner, Art Unit 1631

.....